The Matrix Logarithm: from Theory to Computation

Nick Higham
School of Mathematics
The University of Manchester

higham@ma.man.ac.uk
http://www.ma.man.ac.uk/~higham/

6th European Congress of Mathematics, July 2012
A logarithm of $A \in \mathbb{C}^{n \times n}$ is any matrix $X$ such that $e^X = A$. 

- Implicit definition.
- Properties, classification?
Outline

1. Definition and Properties
2. Applications
3. Theory
4. Computing the Matrix Logarithm and its Fréchet derivative
Matrix algebra developed by Arthur Cayley, FRS (1821–1895) in *Memoir on the Theory of Matrices (1858)*.
- Cayley considered matrix square roots.

Term “matrix” coined in 1850 by James Joseph Sylvester, FRS (1814–1897).
- Gave (1883) first definition of $f(A)$ for general $f$. 
There have been proposed in the literature since 1880 eight distinct definitions of a matric function, by Weyr, Sylvester and Buchheim, Giorgi, Cartan, Fantappiè, Cipolla, Schwerdtfeger and Richter.

— R. F. Rinehart, The Equivalence of Definitions of a Matric Function (1955)
There have been proposed in the literature since 1880 eight distinct definitions of a matric function, by Weyr, Sylvester and Buchheim, Giorgi, Cartan, Fantappiè, Cipolla, Schwerdtfeger and Richter.

— R. F. Rinehart, The Equivalence of Definitions of a Matric Function (1955)
Jordan Canonical Form

\[ Z^{-1}AZ = J = \text{diag}(J_1, \ldots, J_p), \quad J_k = \underbrace{\begin{bmatrix} \lambda_k & 1 \\ & \lambda_k & \ddots \\ & & \ddots & 1 \\ & & & \lambda_k \end{bmatrix}}_{m_k \times m_k} \]

Definition

\[ f(A) = Zf(J)Z^{-1} = Z\text{diag}(f(J_k))Z^{-1}, \]

\[ f(J_k) = \begin{bmatrix} f(\lambda_k) & f'(\lambda_k) & \cdots & \frac{f(m_k-1)(\lambda_k)}{(m_k-1)!} \\ f(\lambda_k) & \ddots & \vdots & \vdots \\ \vdots & \ddots & f'(\lambda_k) \\ f(\lambda_k) & & f(\lambda_k) \end{bmatrix}. \]
Primary and Nonprimary Logarithms

\[ A = \text{diag}(1, 1, e, e). \]

**Primary:** \( \log(A) = \text{diag}(0, 0, 1, 1). \)

**Nonprimary:** \( \log(A) = \text{diag}(0, 2\pi i, 1, 1). \)
Cauchy Integral Theorem

Definition

\[ f(A) = \frac{1}{2\pi i} \int_{\Gamma} f(z)(zl - A)^{-1} \, dz, \]

where \( f \) is analytic on and inside a closed contour \( \Gamma \) that encloses \( \lambda(A) \).
Mercator’s Series

By integrating \((1 + t)^{-1} = 1 - t + t^2 - t^3 + \cdots\) between 0 and \(x\) we obtain Mercator’s series (1668),

\[
\log(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots, \quad |x| < 1.
\]

For \(A \in \mathbb{C}^{n \times n}\),

\[
\log(I + A) = A - \frac{A^2}{2} + \frac{A^3}{3} - \frac{A^4}{4} + \cdots, \quad \rho(A) < 1.
\]
Composite Functions

Theorem

\[ f(t) = g(h(t)) \implies f(A) = g(h(A)). \]

Corollary

\[ \exp(\log(A)) = A \text{ when } \log(A) \text{ is defined.} \]
Composite Functions

**Theorem**

\[ f(t) = g(h(t)) \Rightarrow f(A) = g(h(A)). \]

**Corollary**

\[ \exp(\log(A)) = A \text{ when } \log(A) \text{ is defined.} \]

What about \( \log(\exp(A)) = A \)?

**Matrix unwinding number**

\[ U(A) = \frac{A - \log(\exp(A))}{2\pi i}. \]
Outline

1. Definition and Properties
2. Applications
3. Theory
4. Computing the Matrix Logarithm and its Fréchet derivative
The second-order differential equation

\[
d\frac{d^2 y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y'_0
\]

has solution

\[
y(t) = \cos(\sqrt{A}t)y_0 + (\sqrt{A})^{-1}\sin(\sqrt{A}t)y'_0.
\]
\[
\frac{d^2 y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y_0'
\]

has solution

\[
y(t) = \cos(\sqrt{A}t)y_0 + (\sqrt{A})^{-1} \sin(\sqrt{A}t)y_0'.
\]

But

\[
\begin{bmatrix} y' \\ y \end{bmatrix} = \exp\left( \begin{bmatrix} 0 & -tA \\ tI_n & 0 \end{bmatrix} \right) \begin{bmatrix} y_0' \\ y_0 \end{bmatrix}.
\]
Toolbox of Matrix Functions

\[
\frac{d^2 y}{dt^2} + Ay = 0, \quad y(0) = y_0, \quad y'(0) = y'_0
\]

has solution

\[
y(t) = \cos(\sqrt{A} t) y_0 + (\sqrt{A})^{-1} \sin(\sqrt{A} t) y'_0.
\]

But

\[
\begin{bmatrix}
  y' \\
  y
\end{bmatrix} = \exp \left( \begin{bmatrix}
  0 & -tA \\
  tl_n & 0
\end{bmatrix} \right) \begin{bmatrix}
  y'_0 \\
  y_0
\end{bmatrix}.
\]

- In software want to be able evaluate interesting \( f \) at matrix args as well as scalar args.
- MATLAB has \texttt{expm}, \texttt{logm}, \texttt{sqrtm}, \texttt{funm}. 


University of Manchester
Nick Higham
Matrix Logarithm
13 / 40
Application: Control Theory

Convert **continuous-time system**

\[
\frac{dx}{dt} = Fx(t) + Gu(t),
\]
\[
y = Hx(t) + Ju(t),
\]

to **discrete-time state-space system**

\[
x_{k+1} = Ax_k + Bu_k,
\]
\[
y_k = Hx_k + Ju_k.
\]

Have

\[
A = e^{F\tau}, \quad B = \left( \int_0^\tau e^{Ft} \, dt \right) G,
\]

where \( \tau \) is the sampling period.

MATLAB Control System Toolbox: \texttt{c2d} and \texttt{d2c}.
The Average Eye

First order character of optical system characterized by transference matrix

\[ T = \begin{bmatrix} S & \delta \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{5 \times 5}, \]

where \( S \in \mathbb{R}^{4 \times 4} \) is symplectic:

\[ S^T J S = J = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}. \]

Average \( m^{-1} \sum_{i=1}^{m} T_i \) is not a transference matrix.

Harris (2005) proposes the average \( \exp(m^{-1} \sum_{i=1}^{m} \log(T_i)). \)
Markov Models

- Time-homogeneous continuous-time Markov process with transition probability matrix $P(t) \in \mathbb{R}^{n \times n}$.
- **Transition intensity matrix** $Q$: $q_{ij} \geq 0$ ($i \neq j$),
  $$\sum_{j=1}^{n} q_{ij} = 0, \quad P(t) = e^{Qt}.$$  

For *discrete-time* Markov processes:

**Embeddability problem**

When does a given *stochastic* $P$ have a real logarithm $Q$ that is an *intensity matrix*?
Markov Models (1)—Example

With \( x = -e^{-2\sqrt{3}\pi} \approx -1.9 \times 10^{-5} \),

\[
P = \frac{1}{3} \begin{bmatrix}
1 + 2x & 1 - x & 1 - x \\
1 - x & 1 + 2x & 1 - x \\
1 - x & 1 - x & 1 + 2x
\end{bmatrix}.
\]

- \( P \) diagonalizable, \( \Lambda(P) = \{1, x, x\} \).
- Every primary log complex (can’t have complex conjugate ei’vals).
- Yet a generator is the non-primary log

\[
Q = 2\sqrt{3}\pi \begin{bmatrix}
-2/3 & 1/2 & 1/6 \\
1/6 & -2/3 & 1/2 \\
1/2 & 1/6 & -2/3
\end{bmatrix}.
\]
Suppose $P \equiv P(1)$ has a generator $Q = \log P$. Then $P(t)$ at other times is $P(t) = \exp(Qt)$.

E.g., if $P$ transition matrix for 1 year, 

$$P(1/12) = e^{1/12} \log P \equiv P^{1/12}$$ 

is matrix for 1 month.

**Problem**: $\log P$, $P^{1/k}$ may have wrong sign patterns $\Rightarrow$ “regularize”. 

HIV to Aids Transition

- Estimated 6-month transition matrix.
- Four AIDS-free states and 1 AIDS state.
- 2077 observations (Charitos et al., 2008).

\[
P = \begin{bmatrix}
0.8149 & 0.0738 & 0.0586 & 0.0407 & 0.0120 \\
0.5622 & 0.1752 & 0.1314 & 0.1169 & 0.0143 \\
0.3606 & 0.1860 & 0.1521 & 0.2198 & 0.0815 \\
0.1676 & 0.0636 & 0.1444 & 0.4652 & 0.1592 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}.
\]

Want to estimate the 1-month transition matrix.

\[
\Lambda(P) = \{1, 0.9644, 0.4980, 0.1493, -0.0043\}.
\]

N. J. Higham and L. Lin.
On \( p \)th roots of stochastic matrices, LAA, 2011.
Outline

1. Definition and Properties
2. Applications
3. Theory
4. Computing the Matrix Logarithm and its Fréchet derivative
Logs of $A = I_3$

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 2\pi - 1 & 1 \\ -2\pi & 0 & 0 \\ -2\pi & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 2\pi & 1 \\ -2\pi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$e^B = e^C = e^D = I_3.$$ 

$$\Lambda(C) = \Lambda(D) = \{0, 2\pi i, -2\pi i\}.$$
Theorem (Gantmacher, 1959)

\[ A \in \mathbb{C}^{n \times n} \text{ nonsing with Jordan canonical form } \]
\[ Z^{-1}AZ = J = \text{diag}(J_1, J_2, \ldots, J_p). \text{ All solutions to } e^X = A \]

are given by

\[ X = Z \, \begin{bmatrix} U \\text{diag}(L^{(j_1)}, L^{(j_2)}, \ldots, L^{(j_p)}) \end{bmatrix} U^{-1} Z^{-1}, \]

where

\[ L^{(j_k)} = \log(J_k(\lambda_k)) + 2j_k \pi i l_{m_k}, \]

\[ j_k \in \mathbb{Z} \text{ arbitrary, and } U \text{ an arbitrary nonsing matrix that} \]
\[ \text{commutes with } J. \]
All Solutions of $e^X = A$: Classified

**Theorem**

$A \in \mathbb{C}^{n \times n}$ nonsing: $p$ Jordan blocks, $s$ distinct ei’vals.

$e^X = A$ has a countable infinity of solutions that are **primary functions** of $A$:

$$X_j = Z \text{diag}(L^{(j_1)}, L^{(j_2)}, \ldots, L^{(j_p)})Z^{-1},$$

where $\lambda_i = \lambda_k$ implies $j_i = j_k$. If $s < p$ then $e^X = A$ has **non-primary solutions**

$$X_j(U) = ZU \text{diag}(L^{(j_1)}, L^{(j_2)}, \ldots, L^{(j_p)}) U^{-1} Z^{-1},$$

where $j_k \in \mathbb{Z}$ arbitrary, $U$ arbitrary nonsing with $UJ = JU$, and for each $j$ $\exists$ $i$ and $k$ s.t. $\lambda_i = \lambda_k$ while $j_i \neq j_k$. 


Logs of $A = I_3$ (again)

$$C = \begin{bmatrix} 0 & 2\pi - 1 & 1 \\ -2\pi & 0 & 0 \\ -2\pi & 0 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 & 2\pi & 1 \\ -2\pi & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

$$e^0 = e^C = e^D = I_3. \quad \Lambda(C) = \Lambda(D) = \{0, 2\pi i, -2\pi i\}.$$
Square Roots of Rotations

\[ G(\theta) = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}. \]

\(G(\theta/2)\) is the natural square root of \(G(\theta)\).

For \(\theta = \pi\),

\[ G(\pi) = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad G(\pi/2) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \]

\(G(\pi/2)\) is a nonprimary square root.
Let $A \in \mathbb{C}^{n \times n}$ have no eigenvalues on $\mathbb{R}^-$.

**Principal log**

$X = \log(A)$ denotes unique $X$ such that

- $e^X = A$.
- $-\pi < \text{Im}(\lambda(X)) < \pi$. 

**Principal pth root**

For integer $p > 0$, $X = A^{1/p}$ is unique $X$ such that $X^p = A$.
- $-\pi/p < \text{arg}(\lambda(X)) < \pi/p$. 

---

University of Manchester
Nick Higham
Let \( A \in \mathbb{C}^{n \times n} \) have no eigenvalues on \( \mathbb{R}^- \).

### Principal log

\[ X = \log(A) \] denotes unique \( X \) such that
- \( e^X = A \).
- \( -\pi < \text{Im}(\lambda(X)) < \pi \).

### Principal \( p \)th root

For integer \( p > 0 \), \( X = A^{1/p} \) is unique \( X \) such that
- \( X^p = A \).
- \( -\pi/p < \text{arg}(\lambda(X)) < \pi/p \).
Henry Briggs (1561–1630)

- **Arithmetica Logarithmica** (1624)
- Logarithms to base 10 of 1–20,000 and 90,000–100,000 to **14 decimal places**.
Henry Briggs (1561–1630)

- **Arithmetica Logarithmica** (1624)
  - Logarithms to base 10 of 1–20,000 and 90,000–100,000 to **14 decimal places**.

*Briggs must be viewed as one of the great figures in numerical analysis.*

ARITHMETICA

LOGARITHMICA

SIVE

LOGARITHMORVM

CHILIADES TRIGINTA, PRO

numeris naturali serie crescentibus ab unitate ad

20,000: et a 90,000 ad 100,000. Quorum ope multa

perfeccionem Arithmetica problemata et Geometrica.

HOS NUMEROS PRIMVS

INVENIT CLARISSIMVS VIR IOHANNES

NEPERSVS BARSMERHILTONIJS: eos autem ex eisdem senecta

mutavit, cumque orum et visa illustravit HENRICVS BRIGGSVS,

in celeberrima Academia Oxoniensis Geometrica

professor SAVILLANVS.

DEVS NOBIS VSVRAM VITÆ DEDIT

ET INGENII, TANQUAM PECUNIAE,

NULLA PRESTITVTA DIE.

LONDINI,

Excudebat GVLIELMVS

IONES. 1624.


<table>
<thead>
<tr>
<th>Numbers continuorum Medii inter Denarum &amp; Unitates.</th>
<th>Logarithmi Rationales.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>1.01</td>
<td>0.02</td>
</tr>
<tr>
<td>1.02</td>
<td>0.04</td>
</tr>
<tr>
<td>1.03</td>
<td>0.06</td>
</tr>
<tr>
<td>1.04</td>
<td>0.08</td>
</tr>
<tr>
<td>1.05</td>
<td>0.10</td>
</tr>
<tr>
<td>1.06</td>
<td>0.12</td>
</tr>
<tr>
<td>1.07</td>
<td>0.14</td>
</tr>
<tr>
<td>1.08</td>
<td>0.16</td>
</tr>
<tr>
<td>1.09</td>
<td>0.18</td>
</tr>
<tr>
<td>1.10</td>
<td>0.20</td>
</tr>
<tr>
<td>1.11</td>
<td>0.22</td>
</tr>
<tr>
<td>1.12</td>
<td>0.24</td>
</tr>
<tr>
<td>1.13</td>
<td>0.26</td>
</tr>
<tr>
<td>1.14</td>
<td>0.28</td>
</tr>
<tr>
<td>1.15</td>
<td>0.30</td>
</tr>
<tr>
<td>1.16</td>
<td>0.32</td>
</tr>
<tr>
<td>1.17</td>
<td>0.34</td>
</tr>
<tr>
<td>1.18</td>
<td>0.36</td>
</tr>
<tr>
<td>1.19</td>
<td>0.38</td>
</tr>
<tr>
<td>1.20</td>
<td>0.40</td>
</tr>
<tr>
<td>1.21</td>
<td>0.42</td>
</tr>
<tr>
<td>1.22</td>
<td>0.44</td>
</tr>
<tr>
<td>1.23</td>
<td>0.46</td>
</tr>
<tr>
<td>1.24</td>
<td>0.48</td>
</tr>
<tr>
<td>1.25</td>
<td>0.50</td>
</tr>
<tr>
<td>1.26</td>
<td>0.52</td>
</tr>
<tr>
<td>1.27</td>
<td>0.54</td>
</tr>
<tr>
<td>1.28</td>
<td>0.56</td>
</tr>
<tr>
<td>1.29</td>
<td>0.58</td>
</tr>
<tr>
<td>1.30</td>
<td>0.60</td>
</tr>
<tr>
<td>1.31</td>
<td>0.62</td>
</tr>
<tr>
<td>1.32</td>
<td>0.64</td>
</tr>
<tr>
<td>1.33</td>
<td>0.66</td>
</tr>
<tr>
<td>1.34</td>
<td>0.68</td>
</tr>
<tr>
<td>1.35</td>
<td>0.70</td>
</tr>
<tr>
<td>1.36</td>
<td>0.72</td>
</tr>
<tr>
<td>1.37</td>
<td>0.74</td>
</tr>
<tr>
<td>1.38</td>
<td>0.76</td>
</tr>
<tr>
<td>1.39</td>
<td>0.78</td>
</tr>
<tr>
<td>1.40</td>
<td>0.80</td>
</tr>
<tr>
<td>1.41</td>
<td>0.82</td>
</tr>
<tr>
<td>1.42</td>
<td>0.84</td>
</tr>
<tr>
<td>1.43</td>
<td>0.86</td>
</tr>
<tr>
<td>1.44</td>
<td>0.88</td>
</tr>
<tr>
<td>1.45</td>
<td>0.90</td>
</tr>
<tr>
<td>1.46</td>
<td>0.92</td>
</tr>
<tr>
<td>1.47</td>
<td>0.94</td>
</tr>
<tr>
<td>1.48</td>
<td>0.96</td>
</tr>
<tr>
<td>1.49</td>
<td>0.98</td>
</tr>
<tr>
<td>1.50</td>
<td>1.00</td>
</tr>
</tbody>
</table>

...
Briggs’ Log Method (1617)

\[
\log(ab) = \log a + \log b \quad \Rightarrow \quad \log a = 2 \log a^{1/2}.
\]

Use repeatedly:

\[
\log a = 2^k \log a^{1/2^k}.
\]

Write \( a^{1/2^k} = 1 + x \) and note \( \log(1 + x) \approx x \). Briggs worked to base 10 and used

\[
\log_{10} a \approx 2^k \cdot \log_{10} e \cdot (a^{1/2^k} - 1).
\]
When Does $\log(BC) = \log(B) + \log(C)$?

**Theorem**

Let $B, C \in \mathbb{C}^{n \times n}$ commute and have no ei’vals on $\mathbb{R}^-$. If for every ei’val $\lambda_j$ of $B$ and the corr. ei’val $\mu_j$ of $C$, $|\arg \lambda_j + \arg \mu_j| < \pi$, then $\log(BC) = \log(B) + \log(C)$. 

Proof. $\log(B)$ and $\log(C)$ commute, since $B$ and $C$ do. Therefore $e^{\log(B)} + \log(C) = e^{\log(B)} e^{\log(C)} = BC$. Thus $\log(B) + \log(C)$ is some logarithm of $BC$. Then $\text{Im}(\log(\lambda_j) + \log(\mu_j)) = \arg \lambda_j + \arg \mu_j \in (-\pi, \pi)$, so $\log(B) + \log(C)$ is the principal logarithm of $BC$. 

University of Manchester
Nick Higham
Matrix Logarithm
When Does $\log(BC) = \log(B) + \log(C)$?

**Theorem**

Let $B, C \in \mathbb{C}^{n \times n}$ commute and have no ei’vals on $\mathbb{R}^-$. If for every ei’val $\lambda_j$ of $B$ and the corr. ei’val $\mu_j$ of $C$, $|\arg \lambda_j + \arg \mu_j| < \pi$, then $\log(BC) = \log(B) + \log(C)$.

**Proof.** $\log(B)$ and $\log(C)$ commute, since $B$ and $C$ do. Therefore

$$e^{\log(B)+\log(C)} = e^{\log(B)} e^{\log(C)} = BC.$$

Thus $\log(B) + \log(C)$ is some logarithm of $BC$. Then

$$\text{Im}(\log \lambda_j + \log \mu_j) = \arg \lambda_j + \arg \mu_j \in (-\pi, \pi),$$

so $\log(B) + \log(C)$ is the principal logarithm of $BC$. \qed
Inverse Scaling and Squaring Method

Take $B = C$ in previous theorem:

$$\log A = \log (A^{1/2} \cdot A^{1/2}) = 2 \log (A^{1/2}),$$

since $\arg \lambda (A^{1/2}) \in (-\pi/2, \pi/2)$. 

Use Briggs' idea:

$$\log A = 2^k \log (A^{1/2^k}).$$

Kenney & Laub's (1989) inverse scaling and squaring method:

Bring $A$ close to $I$ by repeated square roots.

Approximate $\log (A^{1/2^s})$ using an\[
\begin{bmatrix} m & m \end{bmatrix}
\end{bmatrix}
\approx \log (1 + x).

Rescale to find $\log (A)$. 

Inverse Scaling and Squaring Method

Take $B = C$ in previous theorem:

$$
\log A = \log(A^{1/2} \cdot A^{1/2}) = 2 \log(A^{1/2}),
$$

since $\arg \lambda(A^{1/2}) \in (-\pi/2, \pi/2)$.

Use Briggs’ idea:

$$
\log A = 2^k \log(A^{1/2^k}).
$$
Inverse Scaling and Squaring Method

Take $B = C$ in previous theorem:

$$\log A = \log \left( A^{1/2} \cdot A^{1/2} \right) = 2 \log(A^{1/2}) ,$$

since $\arg \lambda(A^{1/2}) \in (-\pi/2, \pi/2)$.

Use Briggs’ idea: $\log A = 2^k \log(A^{1/2^k})$.

Kenney & Laub’s (1989) inverse scaling and squaring method:

- Bring $A$ close to $I$ by repeated square roots.
- Approximate $\log(A^{1/2^s})$ using an $[m/m]$ Padé approximant $r_m(x) \approx \log(1 + x)$.
- Rescale to find $\log(A)$. 
Choice of Parameters $s, m$

Must have $\| I - A^{1/2s} \| < 1$.

- Larger Padé degree $m$ means smaller $s$.

Let $h_{2m+1}(X) = e^{r_m(X)} - X - I$.

Assume $\rho(r_m(X)) < \pi$, so $\log(e^{r_m(X)}) = r_m(X)$. Then

$$r_m(X) = \log(I + X + h_{2m+1}(X)) =: \log(I + X + \Delta X),$$

where

$$h_{2m+1}(X) = \sum_{k=2m+1}^{\infty} c_k X^k.$$
Bounding the Backward Error

Want to bound norm of $h_{2m+1}(X) = \sum_{k=2m+1}^{\infty} c_k X^k$.

- Non-normality implies $\rho(A) \ll \|A\|$.  

- Note that

  $$\rho(A) \leq \|A^k\|^{1/k} \leq \|A\|, \quad k = 1: \infty.$$  

and $\lim_{k \to \infty} \|A^k\|^{1/k} = \rho(A)$.

- Use $\|A^k\|^{1/k}$ instead of $\|A\|$ in the truncation bounds.
\[ A = \begin{bmatrix} 0.9 & 500 \\ 0 & -0.5 \end{bmatrix}. \]
Algorithm of Al-Mohy & H (2012)

- Truncation bounds use $\|A^k\|^{1/k}$ rather than $\|A\|$, leading to major benefits in speed and accuracy. Matrix norms not such a blunt tool!
- Use estimates of $\|A^k\|$ (alg of H & Tisseur (2000)).
- Choose $s$ and $m$ to achieve double precision backward error at minimal cost.
- Initial Schur decomposition: $A = QTQ^*$.
- Directly and accurately compute certain elements of $T^{1/2^s} - I$ and $\log(T)$. Use

\[
a^{1/2^s} - 1 = \frac{a - 1}{\prod_{i=1}^{s}(1 + a^{1/2^i})}.
\]
Frechét Derivative of Logarithm

\[ f(A + E) - f(A) - L(A, E) = o(\|E\|). \]

- Integral formula

\[ L(A, E) = \int_0^1 (t(A - I) + I)^{-1} E (t(A - I) + I)^{-1} \, dt. \]

- Method based on

\[ f \left( \begin{bmatrix} X & E \\ 0 & X \end{bmatrix} \right) = \begin{bmatrix} f(X) & L(X, E) \\ 0 & f(X) \end{bmatrix}. \]

- Kenney & Laub (1998): Kronecker–Sylvester alg, Padé of \( \tanh(x)/x \). Requires complex arithmetic.
Algorithm of Al-Mohy, H & Relton (2012)

Fréchet differentiate the ISS algorithm!

1. \( E_0 = E \)
2. for \( i = 1 : s \)
3. Compute \( A^{1/2^i} \).
4. Solve the Sylvester eqn \( A^{1/2^i} E_i + E_i A^{1/2^i} = E_{i-1} \).
5. end
6. \( \log(A) \approx 2^s r_m(A^{1/2^s} - I) \)
7. \( L_{\log}(A, E) \approx 2^s L_{r_m}(A^{1/2^s} - I, E_s) \)

Backward Error Result

\[
\begin{align*}
    r_m(X) &= \log(I + X + \Delta X), \\
    L_{r_m}(X, E) &= L_{\log}(I + X + \Delta X, E + \Delta E).
\end{align*}
\]
Log appears in a growing number of applications.
Have good algorithms for log(A), $L_{\log}(A)$ and estimating the condition number.
If A is real can work entirely in real arithmetic.

- Conditioning of $f(A)$.
- Non-primary functions.
- Functions of structured matrices.
A. H. Al-Mohy and N. J. Higham. 
A new scaling and squaring algorithm for the matrix exponential. 

A. H. Al-Mohy and N. J. Higham. 
Improved inverse scaling and squaring algorithms for the matrix logarithm. 

W. F. Harris.
The average eye.

N. J. Higham.
The Matrix Function Toolbox.
http://www.ma.man.ac.uk/~higham/mftoolbox.

N. J. Higham.
Evaluating Padé approximants of the matrix logarithm.
N. J. Higham.  
*Functions of Matrices: Theory and Computation.*  
xx+425 pp.

N. J. Higham and A. H. Al-Mohy.  
Computing matrix functions.  

N. J. Higham and L. Lin.  
On $p$th roots of stochastic matrices.  
C. S. Kenney and A. J. Laub.
Condition estimates for matrix functions.

C. S. Kenney and A. J. Laub.
A Schur–Fréchet algorithm for computing the logarithm and exponential of a matrix.